The PokéAgent Challenge: Competitive and Long-Context Learning at Scale

Seth Karten* Jake Grigsby† Stephanie Milani‡ Kiran Vodrahalli§ Amy Zhang†
Fei Fang‡ Yuke Zhu† Chi Jin*

Abstract

While frontier AI models excel at language understanding, math reasoning, and code generation, they underperform in out-of-distribution generalization, adaptation to strategic opponents, game-theoretic decision-making, and long-context reasoning and planning. To address these gaps, we introduce the PokéAgent Challenge, leveraging Pokémon's rich multi-agent battle system and expansive role-playing game (RPG) environment. The competition features two complementary tracks: the *Battling Track* evaluates generalization and strategic reasoning under uncertainty in the two-player game of Competitive Pokémon, while the *Speedrunning Track* targets long-horizon planning and decision-making in the Pokémon RPG. Together, our competition tracks unify recent interests in reinforcement learning (RL) and large language model (LLM) research, encouraging collaboration across communities. Pokémon's popularity and internet presence are a key strength of our competition: Participants will have access to a large dataset of over 3.5 million battles and a knowledge base of reference materials and baseline methods. Recent work led by our competition's organizers will provide varied baselines, including rule-based, RL, and LLM-based agents. Our resources will make the PokéAgent challenge accessible while maintaining the complexity needed to drive fundamental advances in decision-making systems.

Keywords Partial Observability, Opponent Modeling, Reinforcement Learning, LLMs, Game AI

1 Competition description

1.1 Background and impact

Progress in artificial general intelligence hinges on benchmarks that resist saturation while demanding coherent integration of diverse capabilities. Traditional benchmarks rapidly approach human performance ceilings, as shown by the Stanford AI Index Report [1], but competitive adversarial environments create sustained open-ended challenges. These domains force agents to continuously adapt strategies in response to evolving opponents, avoiding plateaus until reaching Nash equilibria, which is a property critical for studying superhuman capability takeoff. Simultaneously, the field requires unification of reinforcement learning's trial-and-error grounding with language models' compositional reasoning, particularly for long-horizon tasks requiring thousands of temporally coherent decisions. The ideal benchmark would combine these properties while providing standardized evaluation across methodologies.

*Princeton University, Correspondence: `sethkarten@princeton.edu`
†UT-Austin
‡Carnegie Mellon University
§Google Deepmind

39th Conference on Neural Information Processing Systems (NeurIPS 2025): Competition Track.

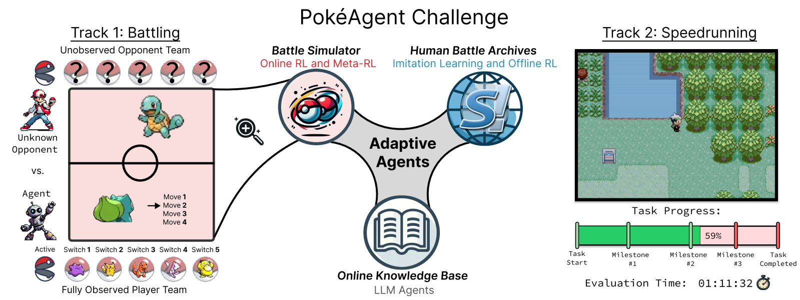

Figure 1: **The PokéAgent Challenge.** Our competition will feature two tracks. **Track 1** utilizes the complex two-player game of Competitive Pokémon to evaluate adaptation and reasoning under uncertainty. **Track 2** deploys agents in the full single-player Pokémon video game, which requires exploration and long-horizon planning. Both tracks benefit from Pokémon's online popularity — we compile datasets to encourage research that incorporates reference materials and demonstrations.

Our PokéAgent Challenge leverages the widespread popularity and complexity of Pokémon to evaluate general decision-making systems. Pokémon is a series of role-playing games (RPGs) in which players explore a large world map to defeat opponents in stochastic turn-based battles. Pokémon emphasizes long-horizon planning and has recently emerged as a testbed for long-context reasoning in language models [2, 3, 4]. While only a subset of the broader RPG gameplay, Pokémon battling has incredible strategic depth and is a popular two-player competitive game in its own right [5]. Competitive battles form a 10^{354}-state partially observable stochastic game where players infer hidden team compositions while executing precise turn-based strategies. The vast space of Pokémon teams creates a more extreme generalization challenge than established research games like chess and poker. This difficulty is compounded by community-driven rule updates that shift strategies and team composition trends, creating non-stationarity that requires robust opponent modeling. Recent work has made significant advances in learning-based competitive Pokémon agents, with approaches ranging from RL to LLM-based planners [6, 7, 8]. However, progress remains fragmented across isolated implementations and evaluations. Standardization is crucial for measuring true advancements in partial observability, opponent modeling, and long-horizon planning.

Broader & Social Impact. This competition advances AI research while addressing real-world challenges. Techniques for opponent modeling have applications in cybersecurity, economics, and finance. Long-horizon planning is an important requirement of autonomous agents in robotics, autonomous driving, and logistics. Importantly, the competition bridges knowledge-based AI (LLMs) and experience-based AI (RL), offering insights into integrating these approaches for robust systems capable of reasoning and adaptation in social domains.

The competition's intersection of reinforcement learning, game theory, planning, and language models is expected to draw significant interest from the NeurIPS community. Pokémon's universal appeal, well-documented mechanics, and competitive scene create an accessible yet deep domain requiring minimal expertise while supporting sophisticated methods. Our competition will provide extensive resources, starter code, and baseline methods of varying difficulty to ensure accessibility. We estimate 300-500 participants based on recent game AI competitions and Pokémon's broad appeal.

1.2 Novelty

The PokéAgent Challenge is an entirely new competition that pioneers Pokémon's dual battle/RPG mechanics as unified benchmarks for decision-making research. While building upon lessons from prior NeurIPS competitions, it introduces three novel elements absent in previous events:

- **Dual-Track Integration:** Unlike single-focus competitions (Neural MMO's multi-agent survival [9] or Lux AI's resource management [10]), we combine competitive partial observability (Track 1) with exploratory long-horizon planning (Track 2). This forces participants to develop methods that balance adversarial reasoning with sustained coherence, which is a requirement absent in existing benchmarks.
- **Community-Driven Non-Stationarity:** Track 1's dynamic rule updates, dictated by the Pokémon player base, create intentional distribution shifts that test robustness to meta-

strategy evolution. This contrasts with static environments like Poker [11] or StarCraft [12], where optimal strategies stabilize post-competition.

- **RPG Speedrunning:** Track 2's speedrunning objective provides concrete metrics (completion time) while preserving Pokémon Emerald's world exploration and teambuilding while following a fairly linear storyline of milestone progression. This bridges the gap between constrained optimization (Chess) and fully open-world benchmarks (MineRL [13, 14]).

Key differentiators from related competitions include:

- **vs Neural MMO:** While Neural MMO tests emergent cooperation, our tracks mandate explicit opponent modeling (Track 1) and procedural route optimization (Track 2).
- **vs MineRL:** Pokémon's turn-based battles enable precise credit assignment and concise evaluation in a fairly linear storyline compared to MineRL's pure open-world, which requires much greater exploration.
- **vs Lux AI:** Lux AI focuses on multi-agent meta-learning and resource management at scale through sequential game sequences with changing dynamics, while both our tracks emphasize partial observability resolution and long-horizon reasoning in addition to the levels of competitive performance found in Lux's fully observable resource optimization tasks.

Though original, the competition leverages proven infrastructure: Track 1 uses the Pokémon Showdown battle simulator (a website with thousands of daily active players), while Track 2 builds on the Pokémon Emerald speedrunning API from our team. All data and code do not overlap with prior competition assets.

1.3 Data

The PokéAgent Challenge provides high-quality data resources for two competition tracks.

For Track 1, we offer a dataset of over 3.5 million Pokémon battles from the Pokémon Showdown platform, including team compositions, movesets, turn-by-turn actions, player ratings, and match results. The data is publicly available, privacy-compliant, and processed to preserve gameplay sequences for supervised and reinforcement learning. Test-time evaluation will occur in live games to ensure out-of-distribution assessment.

For Track 2, we provide a standardized emulation environment for Pokémon Emerald with a Python API offering visual observations, action space definitions, state tracking, frame extraction tools, and utilities for integrating vision-language models. This setup supports advanced speedrunning approaches grounded in real-time processing.

Additionally, participants have access to a comprehensive knowledge base compiled from Bulbapedia [15], featuring over 3,000 articles detailing Pokémon moves, items, abilities, and mechanics to support knowledge-grounded methods. All resources comply with NeurIPS ethical guidelines on privacy and fair use and are designed for accessibility in game AI research.

1.4 Tasks and application scenarios

The PokéAgent Challenge consists of two distinct tracks, each addressing different aspects of AI research:

Track 1: Competitive Battling. Pokémon battles are two-player, stochastic, imperfect-information games with simultaneous actions. Participants construct teams of Pokémon from a combinatorially large design space and compete against an opponent. On each turn, players simultaneously select actions with the goal of depleting the opposing team. Human-level play requires long-term strategic planning across dozens of turns. Turn outcomes are highly stochastic, and gameplay is governed by nuanced mechanics with a long tail of rare but impactful edge cases. Success hinges on more than managing randomness: agents must model their opponent and make decisions under uncertainty. Opponent team details are revealed only when they directly affect the battle, and actions must be chosen without knowledge of the opponent's simultaneous move. The ability to infer hidden team members and predict opponent behavior from prior turns is a defining skill of expert players. The combination of stochastic dynamics, partial observability, and team diversity gives rise to a vast state space that tests an agent's capacity to reason, plan, and generalize. We invite the reader to review Karten et al. [7], Grigsby et al. [8] for a more detailed discussion on the game.

Participants will develop agents capable of winning Pokémon battles according to the official rules of our hosted gamemodes. Gameplay on Pokémon Showdown spans numerous modes and rulesets, each with distinct mechanics and strategies that pose unique challenges for autonomous agents. We will evaluate submissions across three game modes to encourage generalizable methods and discourage brittle, heuristic-based solutions. This track is designed to simulate real-world scenarios where AI systems must rapidly adapt to changing environments and operate effectively with limited prior experience, relying instead on general knowledge and robust reasoning capabilities. Participants may employ self-play reinforcement learning or take advantage of our dataset to study imitation learning or offline RL [16]. The best-of-several-game format of Pokémon tournaments also provides opportunities for meta-learning methods [17] to adapt to their opponents' tendencies over repeat matchups. Alternatively, participants may prefer reasoning-based approaches that do not require explicit task-specific training, including techniques leveraging large language models, retrieval-augmented generation, or test-time planning algorithms.

Track 2: RPG Speedrunning. While battling represents a well-defined competitive environment, the original Pokémon RPGs present an expanded challenge highlighting long-horizon reasoning and exploration. In this track, participants will develop agents that can complete Pokémon Emerald as quickly as possible, which requires agents to navigate the game world, build an effective team, and defeat opponents. To keep evaluation costs and difficulty more manageable, we expect to limit the competition to complete a subset of the game. The task involves long-horizon planning across different game segments, balancing exploration, resource management, and battle strategy, and optimizing for completion time while ensuring reliability. This track simulates scenarios where AI systems must make a series of interconnected decisions to achieve a long-term goal, with each decision affecting future options.

The competition tasks comply with the NeurIPS Code of Ethics. The application scenarios focus on advancing AI research in a widely recognized and engaging domain without harmful consequences. The competition promotes fair play and ethical competition through transparent evaluation criteria and equal access to resources. Players will all have access to baselines methods to help guide their initial solutions for the competition. Additionally, we provide our own hosting of the Pokémon Showdown platform in order to encourage AI-based methods to not saturate the platform used by human players.

1.5 Metrics

In addition to the following metrics, we reserve the right to request full computing infrastructure used for reproduction must be described in full detail, including all code and models.

Track 1: Battle Evaluation. For Track 1 (Pokémon Battles), we will use several complementary metrics to evaluate performance. Our primary metrics will be objective win rates over a large sample of battles against key baselines including rule-based bots, previous state-of-the-art approaches [6, 7, 8], and other participants. The first phase of the competition will also involve established player rating schemes used by the Pokémon community, such as Elo and Glicko-1, with an emphasis on statistically rigorous metrics that account for uncertainty and control for matchmaking bias.

Track 2: Speedrunning Evaluation. For the RPG track, we prioritize metrics that balance completion progress with execution speed. Our primary metric is **Completion Percentage**, measuring progress through a standardized list of critical game milestones. Ties are broken by the time taken to reach the goal state. Completion Time is measured by the in-game timer and is therefore not determined by the runtime of participants' submissions, though we will report the total runtime to evaluate trade-offs between performance and computational cost.

1.6 Baselines, code, and material provided

We will provide comprehensive baseline solutions and starter code for both tracks of the competition.

Track 1. For the Battle Track, we will host a Pokémon Showdown server for players to test their agents. This server will host an array of baseline opponents spanning a range of battling difficulty. Our baselines are grouped into two main categories:

- **Heuristic Baselines.** We create a suite of a dozen heuristic opponents that evaluate core game knowledge. Strategies are based on fundamental Pokémon concepts and re-implementations of opponent policies from official versions of Pokémon. The heuristic set includes random and worse-than-random opponents to help new participants get started. We also include advanced heuristic search engines.
- **Model-Based Baselines.** We include LLM-based baselines such as PokéLLMon [6] (a state-of-the-art LLM-based agent using self-consistency prompting) and PokéChamp [7] (a minimax agent powered by large language models). Finally, we include the Pokémon policies from Metamon [8]. Metamon provides every training checkpoint of 18 different policies trained with various datasets and RL objectives. Metamon, PokéChamp, and PokéLLMon agents have previously been evaluated against human players, which will give participants an interpretable point-of-reference for how their methods are performing without encouraging a large disruption of the public Pokémon Showdown website.

The vast space of possible team configurations makes Pokémon a challenging generalization problem. However, the domain knowledge required to construct high quality teams can pose a significant barrier to entry. To mitigate this, we provide procedurally generated team sets that simulate the choices of top players, based on large-scale public gameplay data. These sets contain thousands of diverse teams, enabling participants to generate varied training distributions and evaluate performance under out-of-distribution (OOD) conditions. We also provide a set of 10–20 curated teams per ruleset. These teams are selected from expert-recommended forums and offers participants safe and effective starting points while reducing evaluation variance in computationally intensive experiments.

Track 2 For the Speedrunning Track, we will provide a comprehensive baseline LLM-based agent approach that demonstrates the feasibility of completing early game milestones. Our baseline agent utilizes a modular architecture with perception, planning, memory, and action components that communicate through a message-passing system. The perception module analyzes game frames using vision-language models to identify the current game state (map navigation, battle, dialogue, or menu interaction). The planning module generates high-level strategies based on the agent's current context, while the memory module maintains a coherent representation of the game state across thousands of timesteps. Finally, the action module translates high-level plans into button presses.

We will also provide robust environment integration tools to streamline development. This includes a Python interface to the Pokémon Emerald emulator built on top of the Stable Retro framework, which offers precise control over game state and rendering. Our starter kit defines standardized observation spaces (RGB frames at 240×160 resolution) and action spaces (the 10 GBA buttons: A, B, START, SELECT, UP, DOWN, LEFT, RIGHT, L, and R).

To help participants track progress, we define concrete milestones throughout the early game that serve as checkpoints for evaluation. These include: (1) reaching Oldale Town, (2) reaching Petalburg City, (3) reaching Route 104, (4) Petalburg Woods, (5) reaching Rustboro City, (6) defeating Roxanne at the first gym, etc. For the competition, we expect agents to progress at least through the third gym (Mauville City/Wattson), with the most sophisticated solutions potentially reaching the fourth or fifth gym.

The repository includes utilities for frame extraction and analysis, allowing participants to save demonstrations of their method. Our web-based visualization tools enable real-time monitoring of agent performance, displaying the current game state, agent observations, planning decisions, and executed actions. This visualization system is particularly valuable for debugging complex agent behaviors across the long time horizons required for effective speedrunning. All code is extensively documented with examples demonstrating how to extend the baseline agent with custom modules or alternative approaches.

All code and materials will be released by June 1, 2025, giving participants ample time to familiarize themselves with the resources before the first leaderboard update. The starter kit will include detailed documentation, tutorials, and examples to help participants get started quickly.

1.7 Website, tutorial and documentation

The competition website will be hosted at `https://pokeagent.github.io` and will serve as the central hub for all competition-related information. The site will feature comprehensive documentation of the competition rules, evaluation criteria, and submission process, along with detailed tutorials on using the provided code and resources. We will post regular updates on the competition timeline and leaderboard, and maintain a FAQ section addressing common questions and issues.

The website will be online within two weeks after acceptance notification and will be regularly updated throughout the competition. Tutorial materials will include starter code, example submissions, a curated list of introductory Pokémon resources, and tips for optimizing agent performance. These materials will be designed to be accessible to participants with varying levels of experience in AI and Pokémon, ensuring that both newcomers and experts can participate effectively.

2 Organizational aspects

2.1 Protocol

Participants can join the competition by registering on the competition website and joining the Discord server.

Track 1 (Pokémon Battles) features a two-phase evaluation process. In Phase 1, we host an open ladder using Pokémon Showdown's matchmaking and rating systems to pair submissions of similar skill levels. The competition server simulates the environment and tracks evaluation metrics, while participants run agents locally and communicate action decisions to the server using the Pokémon Showdown API. This setup reduces computational costs while allowing participants full flexibility to customize or extend the starter code. Teams are free to compete against each other as well as a set of competition-hosted baseline agents. All battles take place in real time on a public website that records evaluation metrics and saves replays for later review. Participants can analyze their matches and monitor their ratings relative to other teams and baselines to assess progress. At the end of Phase 1, teams that achieve a win rate above 50% against a designated set of qualifier baselines will advance to Phase 2. These baselines will be disclosed prior to the Phase 1 deadline, and the qualification threshold may be adjusted based on participation levels. Phase 2 adopts a bracket-style format in which qualifying teams face off in head-to-head matches consisting of multiple battles per pairing, scheduled via Discord. This structure mirrors formal Pokémon Showdown tournament play.

For Track 2 (RPG Speedrunning), we will require full submission of code, three video playthroughs and action input logs. We will use the completion time and game progress of the worst of the 3 playthroughs based on video evidence as an unofficial live leaderboard value. Official leaderboard updates will take place at regular intervals following the schedule in Section 2.3. These results will be determined by replicating the submission in a controlled environment on the organizers' hardware. For additional cheating and overfitting protection, we will implement runtime monitoring to detect unauthorized access to game state and enforce strict time limits per turn.

We will conduct beta tests of our evaluation platform in May-June 2025, with a dry run of the full evaluation pipeline in June 2025 before the official competition launch.

2.2 Rules and Engagement

Eligibility: The competition is open to all, with no restrictions on team size or affiliation.

Tracks: Participants may enter any or all of the two tracks (Battles, Speedrunning), but are encouraged to create a general solution that works for both.

Fair Play: Collusion between teams and multi-accounting in Track 1 ladder rankings are strictly forbidden. Agents must operate within human-observable game states and adhere to 15-second/turn decision limits for battles. Track 2 agents must complete speedruns without crashes or infinite loops and cannot use solution trajectories from human playthroughs. The organizing team reserves final authority to disqualify entries violating these principles.

Transparency: Track 1 winners must provide their detailed methodology or (preferably) code for community benefit. Track 2 submissions require full reproducibility through provided emulator configurations. All participants grant organizers non-exclusive rights to evaluate and showcase their methods for research purposes.

These rules are designed to ensure fair competition while allowing for a wide range of approaches. The focus is on developing effective agents within the constraints of the game environment, rather than exploiting technical loopholes.

Participants will have multiple channels to communicate with organizers and each other:

- **Discord Server:** The primary platform for discussions, questions, and announcements.
- **GitHub Issues:** For reporting bugs or technical issues with the starter code.
- **Competition Website:** Regular updates and announcements will be posted on the website: `https://pokeagent.github.io/`.
- **Dedicated Challenge Email:** `pokeagentchallenge@gmail.com`

2.3 Schedule and readiness

The competition will follow this timeline:

- **May 18, 2025:** Competition acceptance notification
- **May 31, 2025:** Competition website launch with preliminary documentation
- **June 1-June 15, 2025:** Beta testing phase with invited participants, starter kit release
- **June 15, 2025:** Official competition launch
- **July 30, 2025:** First official leaderboard update (both tracks)
- **August 15, 2025:** Second official leaderboard update
- **August 30, 2025:** Third official leaderboard update
- **September 15, 2025:** Final development phase submission deadline
- **September 20-30, 2025:** Tournament phase for Track 1, final evaluations for Track 2
- **October 15, 2025:** Final results announced, winners notified
- **November 15, 2025:** Technical report submission deadline for top teams
- **December 2025:** Presentation at NeurIPS 2025

We have made significant progress in preparing for the PokéAgent Challenge, including collecting a dataset of over 3.5 million Pokémon battles, developing baseline agents, implementing LLM-based agents, creating a data pipeline for RL training, designing evaluation protocols, and establishing the tournament structure. Remaining tasks include finalizing the competition website and documentation, completing the starter kit, setting up the automatic evaluation infrastructure, inviting participants to the Discord server for participant support, and finalizing sponsorships for prizes and computational resources.

2.4 Competition promotion and incentives

To ensure broad participation, we will promote the PokéAgent Challenge through various channels, including academic mailing lists, social media platforms (over 33k combined followers by organizers!), gaming communities, and industry partners. We plan to host introductory webinars on YouTube and Twitch to provide guidance for newcomers. We will encourage participants to stream their evaluation playthroughs live on YouTube and Twitch under a shared tag (# PokéAgent).

Incentives for participation will include monetary prizes (subject to sponsorship), research collaboration opportunities, presentation slots at the NeurIPS 2025 Competition Workshop, computational resources for selected participants, and opportunity to co-author a subsequent NeurIPS 2026 Datasets and benchmark track submission for top solutions if those solutions become a core part of the submission. The exact prize structure will be determined based on secured sponsorships, with tentative amounts ranging from $2,000 to $10,000 for top performers in each track. We have confirmed $50,000 USD from Google Deepmind, and are in talks with a number of companies, including Nunu AI, for additional sponsorship.

We will implement mentorship programs, provide additional resources to teams from institutions with limited infrastructure, and collaborate with organizations focused on underrepresented groups in AI. These efforts aim to lower entry barriers and encourage participation from a wide range of backgrounds and expertise levels.

3 Resources

3.1 Organizing team

Our organizing team brings together expertise in reinforcement learning, game AI, large language models, and competitive Pokémon:

- **Seth Karten** is a Ph.D. candidate at Princeton University and the co-founder of the PokéAgent challenge in charge of curating the data for track 1, providing baseline methods for tracks 1 and 2, and administrating the platform for track 2.
- **Jake Grigsby** is a Ph.D. student at UT Austin and the co-founder of the PokéAgent challenge in charge of curating the data for track 1, providing baseline methods for track 1, and and administrating the platform for track 1.
- **Stephanie Milani** is a Ph.D. candidate at Carnegie Mellon University and co-organizer of the MineRL Diamond and BASALT Competitions at NeurIPS (2019 – 2022). She will ensure robust evaluation protocols and standards and community outreach.
- **Kiran Vodrahalli** is a research scientist at Google DeepMind and a core contributor on Gemini working on long-context reasoning. He will help administer the platform and evaluate for track 2, as well as provide guidance for track 1.
- **Amy Zhang** is an Assistant Professor at UT Austin. She will oversee beta testing and contribute expertise in reinforcement learning.
- **Fei Fang** is an Associate Professor at Carnegie Mellon University. She will oversee beta testing and contribute expertise in multi-agent systems and game-theoretic evaluation.
- **Yuke Zhu** is an Assistant Professor at UT Austin. He will oversee organization and execution of the competition, oversee beta testing for track 1, and contribute expertise in reinforcement learning and language agents.
- **Chi Jin** is an Assistant Professor at Princeton University. He will oversee organization and execution of the competition, oversee beta testing for track 2, and contribute expertise in game theory, multi-agent RL, language agents, and LLM reasoning.

The organizing team represents diversity in academic background, career stage, and research focus, ensuring a well-rounded perspective on competition design and evaluation.

3.2 Resources provided by organizers

The organizing team will provide the following resources:

- **Dataset:** Huggingface-hosted dataset of human battle data and Pokémon game knowledge (Section 1.3).
- **Code:** GitHub repositories containing starter kit code, baseline implementations, and evaluation frameworks (Section 1.6).
- **Infrastructure:** Organizer hosted Pokémon Showdown server for battle evaluations, evaluation servers (over 56 A5000 GPUs, 16 A6000s, and 32 H100s) for speedrunning verification, competition website and leaderboard system on AI Crowd/Kaggle (in talks, tbd).
- **Documentation:** Comprehensive guides, tutorials, and reference materials.
- **Support staff:** Core organizers will provide participants with technical assistance throughout the competition.
- **Sponsors:** We are in talks with several potential sponsors to provide prize money for winners and potentially compute resources for active participants.

3.3 Support requested

We request NeurIPS 2025's support in promoting our competition through its channels, including the website, newsletters, and social media. Additionally, we seek financial assistance for top teams to register and travel to the in-person conference. As this is an in-person event, we will ensure all necessary materials and presentations are prepared, and winners will attend to present their solutions.

A Biography of all team members

Seth Karten is a Ph.D. candidate in the Computer Science Department at Princeton University. His research involves multi-agent learning, reinforcement learning, and foundation models for decision-making. He is the corresponding author of PokéChamp [7], a state-of-the-art approach combining large language models with minimax search for Pokémon battling. Seth has real-world autonomous agent and multi-agent experience from prior work with Waymo and Amazon Robotics. Seth is a recipient of the NSF GRFP fellowship and Princeton's Francis Robbins Upton fellowship.

Jake Grigsby is a Ph.D. student in the Computer Science Department at UT Austin. His research focuses on long-term memory and generalization in reinforcement learning. Most relevant to the proposed competition: he is the author of Metamon [8] — a leading RL effort for Pokémon battling — which will provide many of the datasets and baselines for Track 1 of the PokéAgent Challenge.

Stephanie Milani is a final-year Ph.D. candidate in the Machine Learning Department at Carnegie Mellon University. Her research focuses on building reinforcement learning agents to address human-centered and use-case-inspired challenges. Her research has received best paper awards at the ICML MFM-EAI and NeurIPS GenAI4Health workshops. Stephanie is a 2025 Rising Star in ML & Systems, a 2024 Future Leader in Responsible Data Science & AI, and a 2024 Rising Star in Data Science. She co-organized the MineRL international competition series at NeurIPS.

Kiran Vodrahalli is a research scientist at Google DeepMind, where he is currently a core contributor on Gemini focusing on long-context reasoning. Recently, he has been an avid follower of the Gemini Plays Pokémon Twitch stream. In his spare time, he tinkers with older generation Pokémon games and competitive battling. Previously, he graduated from Columbia with a Ph.D. in Computer Science where he was advised by Daniel Hsu and Alex Andoni, where some of his research was focused on interpretable reinforcement learning and learning in games.

Fei Fang is an Associate Professor at the Software and Societal Systems Department in the School of Computer Science at Carnegie Mellon University. Before joining CMU, she was a Postdoctoral Fellow at the Center for Research on Computation and Society (CRCS) at Harvard University, hosted by David Parkes and Barbara Grosz. She received her Ph.D. from the Department of Computer Science at the University of Southern California advised by Milind Tambe (now at Harvard).

Her research lies in the field of artificial intelligence and multi-agent systems, focusing on integrating machine learning with game theory. Her work has been motivated by and applied to security, sustainability, and mobility domains, contributing to the theme of AI for Social Good. She is the recipient of the Allen Newell Award for Research Excellence 2023, 2022 Sloan Research Fellowship, and IJCAI-21 Computers and Thought Award. She was named to IEEE Intelligent Systems' "AI's 10 to Watch" list for 2020. Her work has won the Best Paper Award at GameSec'23, Deployed Application Award at IAAI'23, Best Paper Honorable Mention at HCOMP'22, Best Paper Runner-Up at AAAI'21, Distinguished Paper at IJCAI-ECAI'18, Innovative Application Award at IAAI'16, the Outstanding Paper Award in Computational Sustainability Track at IJCAI'15. She received an NSF CAREER Award in 2021. Her dissertation is selected as the runner-up for IFAAMAS-16 Victor Lesser Distinguished Dissertation Award, and is selected to be the winner of the William F. Ballhaus, Jr. Prize for Excellence in Graduate Engineering Research as well as the Best Dissertation Award in Computer Science at the University of Southern California.

Amy Zhang is an assistant professor and Texas Instruments/Kilby Fellow in the Department of Electrical and Computer Engineering at UT Austin and an affiliate member of the Texas Robotics Consortium. Her work focuses on improving sample efficiency and generalization of reinforcement learning algorithms through bridging theory and practice, and developing new decision making algorithms for real world problems. Amy completed her PhD in computer science at McGill University and Mila – Quebec Artificial Intelligence Institute, where she was advised by Joelle Pineau and Doina Precup. Previously, she was a research scientist at Facebook AI Research, a postdoctoral fellow at UC Berkeley, and obtained an M.Eng. in EECS and dual B.Sci. degrees in Mathematics and EECS from MIT. She also spent two years on the board of directors for Women in Machine Learning.

Yuke Zhu is an Assistant Professor in the Computer Science Department of UT-Austin, where he directs the Robot Perception and Learning (RPL) Lab. He also co-leads the Generalist Embodied Agent Research (GEAR) lab at NVIDIA Research, which builds foundation models for embodied agents in virtual and physical worlds, particularly for humanoid robots. He focuses on developing intelligent algorithms for generalist robots and embodied agents to reason about and interact with the real world. His research spans robotics, computer vision, and machine learning. He received his Master's and Ph.D. degrees from Stanford University. His work has won various awards and nominations, including the Best Conference Paper Award in ICRA 2019, 2024, the Outstanding Learning Paper Award at ICRA 2022, and the Outstanding Paper Award at NeurIPS 2022. He received the NSF CAREER Award and faculty awards from Amazon, JP Morgan, and Sony Research.

Chi Jin is an Assistant Professor of Electrical and Computer Engineering at Princeton University. He received his Ph.D. in Computer Science from UC Berkeley, advised by Michael I. Jordan. His research focuses on building intelligent agents capable of complex strategy, reasoning, and planning. His group has made important contributions to the mathematical foundations of machine learning, especially in nonconvex optimization, reinforcement learning, and game theory/multi-agent systems. His work has notably advanced the understanding of saddle point escaping in optimization, as well as exploration, function approximation, multi-agent interaction, and partial observability in reinforcement learning. Recently, his work has expanded to enhancing LLM reasoning and developing LLM-based agents for mathematics and games. His group's open-source model, Goedel-prover, achieves state-of-the-art performance in formal reasoning. He is a recipient of the NSF CAREER Award, Sloan Research Fellowship, and Keyes/Emerson Faculty Advancement Award.

References

[1] Nestor Maslej, Loredana Fattorini, Raymond Perrault, Yolanda Gil, Vanessa Parli, Njenga Kariuki, Emily Capstick, Anka Reuel, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Juan Carlos Niebles, Yoav Shoham, Russell Wald, Tobi Walsh, Armin Hamrah, Lapo Santarlasci, Julia Betts Lotufo, Alexandra Rome, Andrew Shi, and Sukrut Oak. The ai index 2025 annual report. Technical report, AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA, April 2025. Available at: `https://hai.stanford.edu/ai-index/2025-ai-index-report`.

[2] Anthropic. Visible extended thinking, 2025. URL `https://www.anthropic.com/research/visible-extended-thinking`.

[3] James Vincent. Debates over ai benchmarking have reached pokémon. TechCrunch, April 2025. URL `https://techcrunch.com/2025/04/14/debates-over-ai-benchmarking-have-reached-pokemon/`. Discusses Google's Gemini model surpassing Anthropic's Claude in Pokémon Red, including details on benchmarking methodology.

[4] Kris Holt. Claude isn't a great pokémon player, and that's okay. Engadget, April 2025. URL `https://www.engadget.com/ai/claude-isnt-a-great-pokemon-player-and-thats-okay-151522448.html`. Analyzes Anthropic's Claude AI's performance and learning process in Pokémon Red, as showcased on Twitch.

[5] KITV Staff. 2024 pokémon world championships in waikiki attracts thousands of attendees. KITV Island News, April 2024. URL `https://www.kitv.com/news/2024-pokemon-world-championships-in-waikiki-attracts-thousands-of-attendees/article_e29b90f6-5c39-11ef-a3fc-73d517846760.html`.

[6] Sihao Hu, Tiansheng Huang, and Ling Liu. Pokéllmon: A human-parity agent for pokémon battles with large language models. arXiv preprint arXiv:2402.01118, 2024.

[7] Seth Karten, Andy Luu Nguyen, and Chi Jin. Pokéchamp: an expert-level minimax language agent. arXiv preprint arXiv:2503.04094, 2025.

[8] Jake Grigsby, Yuqi Xie, Justin Sasek, Steven Zheng, and Yuke Zhu. Human-level competitive pokémon via scalable offline reinforcement learning with transformers, 2025. URL `https://arxiv.org/abs/2504.04395`.

[9] Enhong Liu, Joseph Suarez, Chenhui You, Bo Wu, Bingcheng Chen, Jun Hu, Jiaxin Chen, Xiaolong Zhu, Clare Zhu, Julian Togelius, et al. The neurips 2022 neural mmo challenge: A massively multiagent competition with specialization and trade. arXiv preprint arXiv:2311.03707, 2023.

[10] Stone Tao, Akarsh Kumar, Bovard Doerschuk-Tiberi, Isabelle Pan, Addison Howard, and Hao Su. Lux AI season 3: Multi-agent meta learning at scale. In NeurIPS 2024 Competition Track, 2024. URL `https://openreview.net/forum?id=7t8kWYbOcj`.

[11] Noam Brown and Tuomas Sandholm. Superhuman ai for multiplayer poker. Science, 365 (6456):885–890, 2019.

[12] Oriol Vinyals, Igor Babuschkin, Wojciech M Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H Choi, Richard Powell, Timo Ewalds, Petko Georgiev, et al. Grandmaster level in starcraft ii using multi-agent reinforcement learning. nature, 575(7782):350–354, 2019.

[13] Rohin Shah, Steven H Wang, Cody Wild, Stephanie Milani, Anssi Kanervisto, Vinicius G Goecks, Nicholas Waytowich, David Watkins-Valls, Bharat Prakash, Edmund Mills, et al. Retrospective on the 2021 minerl basalt competition on learning from human feedback. In NeurIPS 2021 Competitions and Demonstrations Track, pages 259–272. PMLR, 2022.

[14] Stephanie Milani, Anssi Kanervisto, Karolis Ramanauskas, Sander Schulhoff, Brandon Houghton, Sharada Mohanty, Byron Galbraith, Ke Chen, Yan Song, Tianze Zhou, Bingquan Yu, He Liu, Kai Guan, Yujing Hu, Tangjie Lv, Federico Malato, Florian Leopold, Amogh Raut, Ville Hautamäki, Andrew Melnik, Shu Ishida, João F. Henriques, Robert Klassert, Walter Laurito, Ellen Novoseller, Vinicius G. Goecks, Nicholas Waytowich, David Watkins, Josh Miller, and Rohin Shah. Towards solving fuzzy tasks with human feedback: A retrospective of the minerl basalt 2022 competition, 2023. URL `https://arxiv.org/abs/2303.13512`.

[15] Bulbapedia wiki pages. `https://bulbapedia.bulbagarden.net/wiki`, 2024.

[16] Sergey Levine, Aviral Kumar, George Tucker, and Justin Fu. Offline reinforcement learning: Tutorial, review, and perspectives on open problems. arXiv preprint arXiv:2005.01643, 2020.

[17] Jacob Beck, Risto Vuorio, Evan Zheran Liu, Zheng Xiong, Luisa Zintgraf, Chelsea Finn, and Shimon Whiteson. A survey of meta-reinforcement learning. arXiv preprint arXiv:2301.08028, 2023.